CN 53-1189/P ISSN 1672-7673

天文数据检索与发布综述*

张海龙[1,2]，冶鑫晨[1]，李慧娟[1]，王 杰[1]，王万琼[1]，托乎提努尔[1]，聂 俊[1]，崔辰州[3]，刘 梁[4]，谌俊毅[5]，陈 肖[6]，薛岩松[6]，何勃亮[3]，李长华[3]，赵 青[8]，肖 健[9]，樊东卫[3]，曹子皇[7]，李珊珊[3]，米琳莹[3]，杨哲睿[4]

(1. 中国科学院新疆天文台，新疆 乌鲁木齐 830011；2. 中国科学院射电天文重点实验室，江苏 南京 210008；
3. 中国科学院国家天文台，北京 100012；4. 中国科学院紫金山天文台，江苏 南京 210008；
5. 中国科学院云南天文台，云南 昆明 650011；6. 中国科学院上海天文台，上海 200030；
7. 中国科学院国家天文台郭守敬望远镜运行与发展中心，北京 100012；
8. 天津科技大学，天津 300222；9. 天津大学，天津 300072)

摘要：海量观测数据及次生数据的高效存储与检索，天文大数据的快速及时处理，加速天文学研究的科学产出等问题，已成为天文观测和天文研究迫切需要解决的难题。以信息技术为支撑的天文大数据的高效分析和处理，帮助天文学家重新审视和了解宇宙。虚拟天文台的出现为全球范围内研究资源的无缝透明连接提供了协议、标准，以协议为基础规范了天文数据的发布与检索方式。以国内外现有的观测设备为基础，综述目前主流天文机构的数据发布与检索相关情况。

关键词：观测数据；数据中心；虚拟天文台；数据归档；数据发布

中图分类号：P111.5；TP391 **文献标识码**：A **文章编号**：1672-7673(2017)02-0212-17

1 介 绍

从古至今，人类对浩瀚宇宙的探索从未停止。步入当代，望远镜制造技术日新月异，主动光学技术、薄镜面拼接技术、自适应光学技术、阵列技术等新技术的出现，使可观测的范围越来越广，效率越来越高，同时观测产生的数据量也大幅增长。随着天文望远镜及终端设备设计与制造技术的不断提高，天文学已从古老的光学观测变为全波段天文学，并且正在进入一个数据雪崩时代[1]。天文信息爆炸式增长不仅给天文学家带来了天文发现的巨大机遇，同时也带来了数据存储、访问和处理等方面的严峻挑战。上百年天体观测资料的积累存在这样一种情况：世界各国的天文资料保存和管理方式各自迥异，不同历史时期的数据保存及管理方式也存在巨大差异。虽然全球范围内已发布的天文数据是自由共享的，但是数据结构的混乱和数据存储的多样化严重影响着天文学家对这些数据资料的提取和使用。

虚拟天文台的出现从根本上解决了全球范围内研究资源的无缝透明连接与访问问题，其目的是以信息技术、海量天文数据为基础建设一个数据密集型网络化天文研究与科普教育平台。目前多国的天文机构以虚拟天文台协议标准为基础实现了天文数据的归档与发布。本文综述天文学在数据密集型时代面临的数据存储与发布问题，分析国内外主要天文机构的数据发布环境，讨论虚拟天文台标准协议和最新相关研究进展。

* 基金项目：国家自然科学基金（U1531125，11503075）；国家重点基础研究发展计划（973计划）（2015CB857100）；中国科学院青年创新促进会；中国科学院天文台站设备更新及重大仪器设备运行专项经费；中国科学院西部之光（XBBS201325）；天文学科技领域云项目（XXH12503-05-05）资助.
收稿日期：2016-06-17；修订日期：2016-07-19
作者简介：张海龙，男，博士. 研究方向：数据密集型研究. Email：zhanghailong@xao.ac.cn

2 中国天文数据库

目前，中国天文数据主要来源于中国科学院下属各天文机构的观测数据。其中，国家天文台数据中心规模最大，数据来源丰富；上海天文台、紫金山天文台、云南天文台和新疆天文台也有自己的天文数据归档与发布环境。

2.1 国家天文台数据中心

国家天文台数据中心[2-3]是中国目前最大的天文数据库，数据库包括国家天文台下属的各天文观测设备产生的天文数据，还有部分其它天文台站的观测数据。国家天文台数据中心的数据查询可以通过中国虚拟天文台(China-VO)网站[4-5]实现。中国虚拟天文台数据中心①的归档数据采用了国际虚拟天文台联盟通用的元数据描述标准，使用 VOTable 格式对数据字段进行描述，符合 UCD1 和 UCD1+标准。星表查询方面，中国虚拟天文台设计了星表椎形检索访问接口，对不同星表数据库搭建了椎形检索服务，实现了分布式天文星表数据的统一查询；天文图像查询方面，中国虚拟天文台设计了天文图像统一的访问接口，查询非常便捷；数据显示方面，中国虚拟天文台目前支持的数据查询输出格式有 HTML、VOTable、CSV、FITS，方便查询和下载。

2.1.1 大天区面积多目标光纤光谱天文望远镜(LAMOST)

大天区面积多目标光纤光谱天文望远镜[6-7]是中国科学院国家天文台的国家级科学研究设备，是反射施密特望远镜，同时可观测 4 000 个左右的目标且拥有 5°视场。LAMOST 的观测数据来自 32 台 CCD(16 台光谱仪，每台光谱仪分红、蓝两端，各 2 台 CCD)采集的原始数据，数据量大约每天 10 GB。对数据按照一定程序进行分析处理后还有更多次生数据产生。

LAMOST 目前正在进行第 4 个观测年计划，2012 年秋季开始第 1 观测年计划。第 1 年的定期光谱测量任务已成功完成，并部分发布，包括光谱测试性测量和常规的周期性测量，数据库已对国内的数据用户和国外的合作伙伴开放共享。

在 DR1② 中包含 220 万条光谱记录，其中信噪比大于 10 的恒星光谱 172 万条，比世界上其他已知巡天项目产生的光谱数总和还要多。DR1 中还有一个 108 万颗恒星的光谱参数星表，是目前全世界最大的恒星光谱参数星表。LAMOST DR1 已经被法国斯特拉斯堡天文数据中心(CDS)的 VizieR 数据库系统收录[8]。

DR1 使用的文件格式是普适图像传输系统(Flexible Image Transport System，FITS)。它是世界各天文台用于数据传输和交换的标准文件格式。FITS 文件命名格式为 spec-MMMMM-YYYY_spXX-FFF.fits，MMMMM 代表本地的儒略日，YYYY 是计划标识符(PLANID)，XX 代表摄谱仪的号码，在 1~16 之间。FFF 代表光纤的编号，在 1 到 250 之间。一个 FITS 文件有一个基本的头部单元、一个可以选择的扩展以及其他可选择的特殊记录。头文件包括强制关键字、文件信息关键字、望远镜参数关键字、观测参数关键字、光谱仪参数关键字、天气条件关键字、数据简化参数关键字、光谱分析结果关键字等。

DR1 主要的数据参数包括：(1)Flux；(2)Inverse Variance；(3)Continuum-Subtracted Flux；(4)Andmask；(5)Ormask。

LAMOST 光谱数据查询系统③实现 5 种功能查询：SQL 查询，矩阵查询，径向查询，位置查询，复杂自定义查询以及光谱显示与光谱缩放等功能。

2.1.2 南极巡天望远镜(AST3)

南极巡天望远镜[9-10]由 3 个改进的 50/68 cm 施密特望远镜组成，配备了 10 K × 10 K STA1600FT

① http://explore.china-vo.org/
② http://dr1.lamost.org/doc/data-production-description
③ http://vision.ia.ac.cn/zh/applications/celestial_spectra/

CCD 相机，安装在南极昆仑站。AST3-1 对 2000 平方度天区进行了超过 3 000 次的 60 s 曝光，其中大麦哲伦星云中心 4 700 多帧、银心区域 3 300 多帧、Wolf Layesing 星 1 000 多帧，这些数据由国家天文台的南极天文团队处理。

2.1.3 Beijing-Arizona 巡天项目(BASS)

BASS④,⑤是一个大视场多色域、对北银冠 5 000 平方度区域进行巡天的观测项目。该项目使用 Kitt Peak 天文台的 2.3 m Bok 反射望远镜，由国家天文台和亚利桑那大学合作进行为期 3 到 4 年的观测，巡天数据将用于确定 DESI 暗能量光谱仪巡天项目的有效搜寻目标。目前此项目还在进行中，2015 年 12 月 25 日 BASS EDR 数据在内部发布，DR1 预计在 2016 年夏季发布，DR2 预计在 2016 年冬季发布。

2.1.4 中国之星小望远镜阵(CSTAR)

第 24 次中国南极科考队在南极昆仑站部署了中国之星小望远镜阵⑥，它由 4 个 14.5 cm 的光学望远镜组成，每一个都有不同的滤波器。在 158 天 1 728 小时的观测中，每个望远镜每隔 30 s 记录一张图像。CSTAR #1 记录了超过 30 万张图像，每张图有超过一万个亮于 16 等的目标被识别，通过对比，可以从这些数据中发现超新星、伽马射线暴、小行星等。

2.1.5 南银冠 U 波段巡天(SCUSS)

南银冠 U 波段巡天[11]是中国科学院国家天文台和 Steward 天文台的一个国际合作项目。该项目使用 Kitt Peak 天文台的 2.3 m Bok 反射望远镜，在 U 波段对南银极 3 700 平方度进行观测。这台望远镜配备了 4 K × 4 K 的 CCD，观测视场是 1.031 6°×1.031 6°，极限观测星等为 23 等，从 2010 年 9 月到 2012 年 12 月进行观测。项目主要为 LAMOST 提供 U 波段星表，合并 SDSS III 的数据后可以得到南银极从 355 nm 到 847 nm 的光谱能量分布。目前南银冠 U 波段巡天可以查询和下载的数据包括⑦,⑧：星表数据、影像数据、天体自行数据。

2.1.6 怀柔太阳观测站(Huairou SOS)

怀柔太阳观测站[12]建于 1984 年，是著名的太阳磁场观测基地。怀柔观测站最主要的观测设备由 5 台不同功能的望远镜组成，望远镜统一组装在带有光电引导的跟踪设备上。5 台望远镜分别是：35 cm 太阳磁场望远镜、10 cm 全日面矢量磁场和视向速度场望远镜、14 cm 色球望远镜、8 cm 全日面单色像望远镜和 60 cm 多通道太阳望远镜。主要用于太阳物理基础研究、日地关系应用基础研究以及太阳活动对空间环境和通讯干扰预报等应用研究，怀柔太阳站观测数据可在中国虚拟天文台数据中心查询。

2.1.7 太阳射电宽带动态频谱仪(SBRS)

太阳射电宽带动态频谱仪⑨主要观测太阳大气无线电辐射动态频谱随时间的变化特征。它有 5 个频谱仪和 1 个单频接收机，观测数据可使用 IDL 语言读取。目前数据库中可以查询的数据为：1994 年至 2002 年 1.0 GHz~2.0 GHz 频带数据；2002 年至 2006 年以及 2009 年 1.10 GHz~2.06 GHz 频带数据；1996 年至今 2.6 GHz~3.8 GHz 频带数据；1999 年至今 5.2 GHz~7.6 GHz 频带数据；1977 年至 2007 年以及 2011 年至今 2 840 MHz ± 5 MHz 单频带数据。

2.2 紫金山天文台

紫金山天文台对外开放的数据库⑩包括毫米波射电天文数据库、太阳射电频谱观测数据库、近地天体望远镜数据库、太阳光谱数据库，各数据库已实现在线访问。紫金山天文台已经归档的数据总量

④ http://batc.bao.ac.cn/BASS/doku.php
⑤ http://explore.china-vo.org/data/bass/intro
⑥ http://explore.china-vo.org/data/cstar/intro
⑦ http://batc.bao.ac.cn/BASS/doku.php? id=scuss:scuss
⑧ http://batc.bao.ac.cn/BASS/doku.php? id=scuss:scuss
⑨ http://explore.china-vo.org/data/sbrs/intro
⑩ http://www.pmo.ac.cn/qt/twsjk/

超过 45 TB，项目子库和总站均提供网络式数据查询服务，这些数据库可以分别独立查询，其中 2000 ~2014 年的数据可以整合查询。在页面输入参数后可返回数据列表，可以一键下载所有数据或选择下载所需数据。当前紫金山天文台正在逐步实现具有特色的天文数据库，实现多源天文数据的一站、一键式共享服务模式。

2.2.1　毫米波射电天文数据库

毫米波射电天文数据库[13]是针对紫金山天文台 13.7 毫米波射电望远镜近十年观测获取的谱线数据研发的，13.7 毫米波射电望远镜是目前为止我国唯一工作在毫米波段的天文观测设备。该数据库 2010 年 5 月投入使用，2010 年底 13.7 毫米波射电望远镜安装了超导成像频谱仪，开始对银河系平面 $\pm 5^{\circ}$ 范围内 CO 同位素 3 条分子谱线进行大尺度巡天，数据量随之显著上升，可以通过坐标、半径或观测目标查询，查询结果可以是单点谱线或飞行观测模式(On The Fly，OTF)成图谱线。

毫米波射电天文数据库的建设，结合了天文学领域重大科学工程，形成天文观测专题数据库并包含最新的观测资料。目前该数据库已对外开放，提供数据资源共享服务。迄今为止已有国内科研单位的天文学家以及美、日、德、英、俄等十多个国家的研究人员利用该望远镜的观测数据研究恒星形成、宇宙的起源及演化，并取得了一大批研究成果。

2.2.2　太阳射电频谱天文数据库

太阳射电频谱天文数据库⑪包含了紫金山天文台近 15 年观测获取的太阳射频数据，太阳射电频谱仪每 5 ms 采样一次，工作带宽为 10 MHz，可对工作范围内 300 个波段进行观测。该数据库从 2014 年底投入试运行，观测数据分为原始数据和频谱图片，频谱图片是处理得到的反映太阳频谱和射电强度对应关系的数据，获取方式可以通过年份列表获取，也可以通过搜索日期获取原始数据或者频谱图片。

2.2.3　近地天体望远镜数据库

近地天体望远镜数据库⑫存储紫金山天文台盱眙观测站的近地天体望远镜拍摄的图像数据。近地天体望远镜采用施密特光学系统，主镜和改正镜的口径分别为 120 cm 和 104 cm，焦距 180 cm。观测终端系统由两台 CCD 组成，分别使用 4 K × 4 K CCD 相机和 10 K × 10 K CCD 相机，单帧覆盖视场 3.78 平方度和 9 平方度。数据格式为 FITS，可通过数据表的方式查看与下载，也可以通过观测起始时间、中心坐标查询。目前可公开查询的数据截止到 2013 年 12 月 31 日，最新的数据需要验证登录后才能查询。

2.2.4　太阳光谱数据库

太阳光谱数据库⑬的数据包含紫金山天文台自 2001 年以来观测的太阳光谱数据，观测仪器为多通道近红外太阳光谱仪，它可在 Hα、CaII、HeI 3 个波段观测。太阳光谱数据库可通过日期、起始时间以及波段等参数查询，另外该数据库还支持查询耀斑的发生日期。

2.3　上海天文台

上海天文台拥有甚长基线干涉测量(Very Long Baseline Interferometry，VLBI)观测站、国际 VLBI 网数据处理中心、1.56 m 光学望远镜、60 cm 卫星激光测距望远镜等多项现代空间天文观测设备。目前上海天文台开放的天文数据库有 1.56 m 天体测量望远镜科学数据库以及 65 m 射电望远镜科学数据库。

2.3.1　天体测量数据库

1.56 m 光学望远镜⑭位于上海佘山，海拔 97 m，是中国科学院光学联合开放实验室的主要观测设备之一。1989 年启用，主要探测设备是美国 Photometrics 公司的 CCD 照相机，芯片大小为 1 024 × 1 024 像素，视场为 4′17″，分辨率为 0.25 像素，并配有 B、V、R、I 宽带滤光片。目前上海

⑪　http://www.pmo.csdb.cn/index1.php
⑫　http://www.cneost.org/index.html
⑬　http://www.pmo.csdb.cn/index1.php
⑭　http://center.shao.ac.cn/156cm.htm

1.56 m 光学望远镜的研究领域主要包括 Blazar 光变的研究、活动星系核的国际联测、星团的运动学和动力学研究、球状星团小变幅新类型变星、类星体短时标光变的探测、土星和天王星卫星定位观测、射电源光学对应体的精确测定、激变变星的观测以及太阳系小天体的观测等。

上海天文台天体测量数据库的建设是国家科技基础条件平台建设项目基础科学数据共享网——理化天文空间生物的课题之一，目前观测申请者拥有数据的优先使用权，数据发布一年之内其他用户没有使用数据的权限，数据的有效保护期为 1 年，1 年之后任何普通用户均有权使用 1.56 m 望远镜观测的天体测量数据。

2.3.2 射电数据库

上海天文台 65 m 射电望远镜⑮是亚洲最大的全向可转动射电望远镜[14]，目前总体性能在国际同类型望远镜中位列前 4 名。65 m 射电望远镜具有主动面调整系统、8 个波段的双极化接收机(L、S、C、X、Ku、K、Ka 和 Q 波段)、中国数字采集终端(Chinese VLBI Data Acquisition System，CDAS)、数字基带转换器(Dgital Basic Band Converter，DBBC)、数字终端系统(Digital Backend System，DIBAS)和连续谱终端、氢原子钟和时频比对设备等。65 m 射电望远镜的数据属于上海天文台射电天文科学与技术研究室，尚未对外发布。

2.3.3 VLBI 射电天文和深空测量数据库

甚长基线干涉测量是 20 世纪 60 年代后期发展起来的射电天文新技术。由于它的高测量精度和高分辨率，成为当代天文学和地球动力学研究的重要工具；同时，甚长基线干涉测量技术也在空间探测中对航天器的高精度测定轨发挥了重要作用。

中国 VLBI 网由上海佘山站、北京密云站、云南昆明站、新疆南山站以及上海 VLBI 数据处理中心组成。中国 VLBI 网在射电天文和深空测量等方面开展了 VLBI 观测：在国家探月工程嫦娥一号和嫦娥二号任务中，进行测定轨测量观测，为卫星测控的顺利实施发挥了关键作用；在陆态网任务中，开展了 VLBI 大地测量观测，取得了高精度的台站位置数据；在天体物理方面，开展了脉冲星、活动星系核等观测。

VLBI 射电天文和深空测量数据库系统包括 5 个子库⑯：

(1) VLBI 观测计划数据库。主要负责存放用户信息，发布观测任务通告，接收和评估科学用户的观测申请书；保存所有批准的观测代码及其观测进展，参与观测台站和观测源等信息。

(2) VLBI 深空测量数据库。提供嫦娥一号卫星、嫦娥二号卫星、金星快车、火星快车等深空测量观测过程中的任务书、观测纲要、误差修正、相关处理、相关后处理、定轨定位和运行管理结果文件。

(3) VLBI 天体物理数据库。提供中国 VLBI 网开展的天体物理观测过程中的任务书、观测纲要、误差修正、相关处理、FITSIDI 和处理结果文件。

(4) VLBI 天体测量和大地测量数据库。提供陆态网 VLBI 观测和一些天体测量观测过程中的任务书、观测纲要、误差修正、相关处理、相关后处理、NGS 和处理结果文件。

(5) VLBI 文献数据库。主要提供关于 VLBI 观测和有关数据应用方面的论文和书籍等信息。

VLBI 射电天文和深空测量数据库，目前已经收集了使用中国 VLBI 网开展的射电天文(天体物理、天体测量和大地测量)科学数据，以及深空探测任务(嫦娥一号和嫦娥二号任务)的测量数据，总数据量达到 3 TB。所有数据通过一定授权后，供科学家使用。

2.4 新疆天文台

新疆天文台数据中心可以通过两个域名访问⑰,⑱，两个域名分别指向两台互为备份的数据存储服

⑮ http://65m.shao.cas.cn/
⑯ http://119.78.226.68
⑰ http://www.pmo.ac.cn/qt/twsjk/
⑱ http://vo.xao.ac.cn

务器，两台数据发布设备一台位于新疆天文台本部，一台位于新疆天文台南山观测站。在虚拟天文台标准与协议的框架下，新疆天文台数据中心已能提供多方面的数据检索服务。

2.4.1　天文数据查询语言

天文数据查询语言(Astronomical Data Query Language，ADQL⑲)是一种类似 SQL 的语言，用来检索天文星表或者其他以表形式存在的数据集。通过天文数据查询语言主页面⑳可实现对新疆天文台已发布的、允许使用的所有数据进行查询操作。

2.4.2　锥形检索

锥形检索㉑是在指定半径的一个锥形天区中实现的一种检索。新疆天文台发布了 PPMXL 星表数据锥形检索服务㉒，各课题组数据及其他星表数据均支持锥形检索。通过并行计算技术及伪球面天球分区技术使检索速度得到提高。以 PPMXL 星表数据为例，10 亿目标中检索单条数据返回时间在毫秒量级。

2.4.3　表访问协议

表访问协议(TAP㉓) 定义了一个服务协议访问表格元数据信息，包括天文目录以及通用数据库表，提供访问数据库和表的元数据以及实际的表数据。新疆天文台发布了表访问协议查询服务页面㉔，服务可以实现已注册到国际虚拟天文台联盟(International Virtual Observatory Alliance，IVOA)中的表及字段数据的元数据信息查询。

2.4.4　简单应用程序消息传递协议

简单应用程序消息传递协议㉕(Simple Application Messaging Protocol，SAMP)可以实现天文软件间的互操作和数据交换。新疆天文台发布的数据服务均支持简单应用程序消息传递协议，可将查询结果直接传递给标准虚拟天文台工具，进行数据挖掘、可视化等相关操作。

2.4.5　新疆天文台在线交叉证认服务

天文交叉认证是实现多波段数据融合的基础，是加深对天体的认识，促进天文学新发展的关键一步，同时也是大型望远镜在观测选源阶段进行星体类型识别的核心技术。

新疆天文台发布的在线交叉认证服务㉖目前可以实现远程 URL 方式及本地文件上传带有 UCD 信息的 RA(ucd：pos.eq.ra)、DEC(ucd：pos.eq.dec)的 VOTable 格式星表数据文件。选择目标星表进行交叉认证，可以是新疆天文台已发布数据中的任何一张表，确定所要认证的星表后，给定所要匹配的搜索半径(误差半径)，由于单台观测设备的最高分辨能力及误差半径不同，所以根据星表数据产生设备的具体指标输入搜索半径，之后点击 GO 按钮，可以得到认证结果。

2.4.6　新疆天文台观测数据服务

新疆天文台数据中心已基本完成对南山 25 m 射电望远镜观测产生的脉冲星、IDV、分子谱线等数据的在线存储、归档、备份，并向全台相关研究人员提供数据的存储与发布等服务。以脉冲星数据为例，新疆天文台建立了脉冲星数据检索平台㉗，脉冲星数据查询实现了锥形检索与基于元数据信息的多目标约束检索。

2.4.7　检索结果输出格式

新疆天文台数据中心支持多种数据格式输出，可选择 HTML、FITS Table、CSV、JSON、VOTable、

---

⑲ http://www.ivoa.net/documents/latest/ADQL.html
⑳ http://data.xao.ac.cn/_system_/adql/query/form/
㉑ http://www.ivoa.net/documents/latest/ConeSearch.html
㉒ http://data.xao.ac.cn/ppmxl/q/cone/form
㉓ http://www.ivoa.net/documents/TAP/
㉔ http://data.xao.ac.cn/glots/q/plain/form
㉕ http://www.ivoa.net/documents/SAMP/
㉖ http://data.xao.ac.cn/cross/q/match/form
㉗ http://data.xao.ac.cn/pul/pulsar/q/form

tar 等形式的数据输出，如果数据量大且文件数较多时可选择 tar 格式，直接打包下载所有满足条件的数据。针对结果返回数据量有可能过大的情况，设置了 Limit to 字段，约束结果输出的行数。由于浏览器响应需要时间，数据量过大导致其无响应，可根据需要调整相关数值。

2.5　云南天文台

云南天文台主要观测设备有㉘ 20 世纪 80 年代由德国引进的 1 m 光学望远镜，1.2 m 国产地平式光学望远镜，2006 年从英国引进的 2.4 m 光学望远镜(安装在丽江天文观测站)，抚仙湖太阳观测站新建的 1 m 真空太阳望远镜在 2015 年 5 月通过鉴定并投入使用。除了以上观测设备外，云南天文台还拥有 1 台主要用于探月的 40 m 射电望远镜。目前云南天文台丽江 2.4 m 光学望远镜和 1 m 真空太阳望远镜的观测数据已对外开放。

2.5.1　丽江 2.4 m 望远镜

丽江 2.4 m 望远镜是一架地平式反射光学望远镜，隶属中国科学院云南天文台丽江天文观测站，是东南亚地区最大的通用型光学望远镜，位于东经 100°01′51″、北纬 26°42′32″，海拔 3 193 m。该望远镜得益于良好的观测条件，可观测的极限星等达到 25 等。观测目标主要包括太阳系外行星、银河系内恒星、邻近星系及其恒星和星团、遥远类星体、伽马射线暴、宇宙大尺度结构等。

2.4 m 望远镜于 2008 年 5 月建成使用，2015 年底已经积累观测数据数十 TB。目前 2.4 m 望远镜的数据已经并入中国虚拟天文台数据库，可以直接在中国虚拟天文台的丽江 2.4 m 望远镜数据库㉙中查询。

2.5.2　1 m 真空太阳望远镜

1 m 真空太阳望远镜是我国太阳物理和空间科学对太阳进行光学和近红外观测的主力设备，其主要科学目标是在 0.3~2.5 um 波段对太阳进行高分辨率成像和光谱观测。望远镜于 2010 年 9 月开始试观测，2012 年 9 月正式投入运行。目前获得超过 500 TB 的太阳活动高分辨率观测数据，并于 2015 年 5 月 14 日通过成果鉴定。从 2012 年开始观测至今，1 m 真空太阳望远镜收集了超过 2 000 万个 FITS 文件，每日最大观测记录为 12 万个文件。为了存储和快速查询这些数据，使用 Fastbit 的非结构化数据库(NoSQL)[15]，在 4 000 万条记录的数据库检索 Fastbit 的数据库相对速度较快，完全可以满足 1 m 太阳望远镜的数据存储和查询需求。1 m 太阳望远镜每日观测的影像数据可以在抚仙湖观测站官网上查询，并可下载连续观测的视频资料。数据资料的查询可以按照观测日期查询，也可以通过指定观测日期范围、观测坐标、观测波段等指标查询。其他更为详尽的数据，可以通过电子邮件㉚的方式申请获取。

3　斯特拉斯堡天文数据中心

斯特拉斯堡天文数据中心㉛(Strasbourg Astronomical Data Center, CDS)创建于 1972 年，当时称为恒星数据中心，1983 年改为斯特拉斯堡天文数据中心。斯特拉斯堡天文数据中心在传播有价值的科学数据方面一直处于领先位置。目前致力于收集并向世界各地发布天文数据以及相关信息。

3.1　斯特拉斯堡天文数据中心数据收集

斯特拉斯堡天文数据中心目前的工作重心是：(1)收集基于 VizieR 形式存储的天体信息；(2)通过严格的评议和比较更新现有信息；(3)将分析后的数据发布给相关天文机构；(4)将这些数据用于研究工作。

数据中心成立以来，天文学家、档案文献学家以及信息工程专业的工程师共同应对数据量的不断增加和数据复杂程度的增加带来的问题[16]。天文学家利用他们对项目的理解和用户的需求确保数据

㉘　http://fso.ynao.ac.cn/
㉙　http://explore.china-vo.org/data/gmg240/f
㉚　http://nvst_data@ynao.ac.cn
㉛　http://cdsweb.u-strasbg.fr/

的准确性和相关性。计算机工程师通过维护数据框架和添加有用的工具提高检索和重用数据的效率。档案文献学家根据数据内容管理数据。通过 1971 年以来统计的员工数据分析，这 3 种职位在数据中心同步增加，人数基本保持同步。

3.2 斯特拉斯堡天文数据中心的数据服务

斯特拉斯堡天文数据中心目前拥有 3 项服务㉜：(1) SIMBAD 天文数据库，这是一个可供全世界参考的天文目标识别数据库；(2) VizieR 数据库，是目前世界上最大的在线天文星表数据库，收集了已经发表的天文星表和图表；(3) Aladin 互动天图软件，用于存取、可视化分析天文图像、观测、星表、数据库及其他相关数据。

3.2.1 SIMBAD 天文数据库

SIMBAD 数据库㉝提供了太阳系外天体的基本数据、测量、参考文献等数据。在数据库中可以直接查询目标的名称，也可以使用坐标等其他方法查找。截止目前，数据库收录了 8 041 846 个目标数据，22 617 406 个标识，313 610 篇参考论文。SIMBAD 还发布了 SIMWATCH 工具，用户可以关注在 SIMBAD 数据库中记录的自己感兴趣的天体目标，当有关此天体的新论文发表时 SIMWATCH 工具会通知用户查看。

3.2.2 VizieR 星表数据库

VizieR㉞ 提供了目前最完整的正式发表的天文星表和数据列表，这些数据被组织在一个自我归档的数据库中。VizieR 提供的查询工具允许用户选择相关的数据列表并用标准格式导出，目前有 14 251 个星表可以查询使用，而且这个数字在不断增加。星表的查询可以通过不同波段、不同观测项目、不同天文学领域进行筛选，这些数据基本上覆盖了全波段、绝大部分公开的天文观测项目和所有的天文学领域。VizieR 提供了 3 个工具：Photometry viewer，用于查询 VizieR 中所有星表中的光度信息；TAP VizieR，一个可以使用 SQL 语句查询数据库元数据信息的工具；CDS cross-match service，用于快速对两张表之间的数据进行交叉比对，这些表的数据可以是 VizieR 中的，也可以是 SIMBAD 中的。VizieR 星表和光谱数据均可以通过 IVOA 的标准接口和检索方式访问，可以方便地查询或者交叉使用。

3.2.3 Aladin 天图软件

Aladin㉟ 可以交互式查看天文数字图像，从天文星表和数据库中获取星体的条目，交互式访问 Simbad 数据库、VizieR 服务以及其他已知天文项目中的相关数据和信息。

4 欧洲南方天文台科学数据中心

欧洲南方天文台科学数据中心㊱,㊲是世界上最大的天文学数据归档中心，科学数据中心长期、安全地存储欧洲南方天文台所属的天文数据，并且这些数据向全世界的研究人员开放。数据库中有原始数据，也有经过不同程度处理的数据，所有信息可以通过欧南台的页面㊳进行查询。

4.1 原始数据概况

当前欧南台的观测数据主要来自智利 Paranal 天文台的甚大望远镜(LVT)以及辅助望远镜和干涉仪、可见光和红外巡天望远镜(VISTA)、LVT 巡天望远镜(VST)等。欧南台科学数据中心的原始数据分为成像数据、光谱测量数据、干涉度量数据。在欧南台数据中心，可以根据需求查询所需要的仪器数据，具体分类见表 1。

㉜ http://cdsweb.u-strasbg.fr/about
㉝ http://simbad.u-strasbg.fr/simbad/
㉞ http://vizier.u-strasbg.fr/
㉟ http://aladin.u-strasbg.fr/
㊱ http://archive.eso.org/cms.html
㊲ http://archive.eso.org/cms/tools-documentation/dicb/dicd_v5.pdf(DataInterfaceControl Document)
㊳ http://archive.eso.org/

表 1 观测模式与波段对应仪器表

Table 1 Instrument Forms with Observation Mode and Waveband

Observation Mode	UV	Optical	Near/Mid-IR	Submm
Imaging		FORS1 FORS2 SPHERE VIMOS VST/OmegaCAM	HAWKI ISAAC NACOSPHEREVISIR VISTA/VIRCAM	APEX
Spectroscopy	GIRAFFE UVES X-Shooter	FORS1 FORS2 GIRAFFE MUSE SPHERE UVES VIMOS X-Shooter	CRIRES ISAAC KMOS NACO SINFONISPHEREVISIR X-Shooter	APEX
Polarimetry		FORS1 FORS2 SPHERE	ISAAC NACO	
Interferometry			AMBER MIDI PIONIER	
Other		FORS2	NACO	APEX

4.2 处理后的数据概况

欧洲南方天文台每年新增的仪器可以在科学数据中心查询相关信息。科学数据中心除了归档仪器产生的原始数据外，还提供处理过的数据产品的归档服务[17]。

ESO phase 3㊴提供了简化、标准化的数据产品，所有的数据使用标准格式存储，其中包括欧南台自身设备产生数据的处理后数据，也有合作项目的处理数据。这些处理包括使用延展的 FITS 文件存储 VISTA/VIRCAM 和 VST/OmegaCam 的不连续图像，同时对图像做了形变和光度修正；将 VLT 和 NTT 的一维光谱数据校准为二进制光谱 FITS 文件，每一个数据组被存放在包含一个二进制表的独立单元内，目前还有 UVES、XSHOOTER、HARPS 和 FLAMES/GIRAFFE 处理后数据，对于新观测的光谱数据会自动处理，之后立刻发布。ESO phase 3 还有部分处理后的星表，可以通过查询界面查询单个或者多个来源的星表数据，目标查询界面只有授权用户可以访问。

FEROS & HARPS 数据库提供了处理过的 FEROS 数据(截止 2005 年 4 月)，处理的 HARPS 光谱仪的数据(截止 2003 年 10 月)以及 HARPS 的偏振测定数据(截止 2010 年 10 月)。

其他处理后的数据产品包括：(1)年轻星团的凌日行星和褐矮星的数据集；(2)科罗系外行星探测器(Convection Rotation and planetary Transits，COROT)空间星震观测数据集；(3)对广角成像仪(Wide Field Imager，WFI)的 U、B、V、R、I 波段的光学成像数据进行处理的数据集 GaBoDS/WFI，是欧南台 Deep Public Survey(DPS)的一部分。

除了上面的内容，在欧南台科学数据中心包括：APEX(Atacama Pathfinder Experiment)亚毫米波段原始数据，以及 Atacama 项目的信息存档；欧南台观测计划查询表，可以查询观测项目信息和调度信息。科学数据中心的数据还在不断更新中，部分处理过的星表也可以通过 CDS 实现查询。

5 美国宇航局空间科学数据中心

美国宇航局空间科学数据中心负责对美国宇航局太空科学任务获取的数据进行归档，并将这些数据永久保存。它提供在线访问页面㊵。数据中心主要存档数据来自太空飞船以及空间科学实验数据，除存储数据外还提供数据管理标准与技术支持。目前该中心把数据分学科归档，包括天体物理、太阳物理学、太阳系探索、图像资源等。

5.1 观测设备及数据管理

5.1.1 高能天体物理科学存档中心

高能天体物理科学存档中心㊶是美国宇航局研究黑洞、宇宙大爆炸等高能宇宙现象电磁辐射的主

㊴ http://www.eso.org/sci/observing/phase3/p3sdpstd.pdf(ESO Science Data Products Standard)

㊵ http://nssdc.gsfc.nasa.gov/

㊶ http://heasarc.gsfc.nasa.gov/

要数据存档机构。中心于 2008 年合并了微波背景辐射研究机构，目前包含的高能数据有：远紫外、X 射线、伽马射线和从太空任务、探测气球、地面设施研究中获取的亚毫米、毫米以及厘米波段的宇宙微波背景辐射数据。

在科学存档中心创立之初最大的问题是各任务数据格式不统一，同时软件不能重用。为了解决这个问题，科学存档中心采用了单一的 FITS 标准数据格式存放数据[18-19]，同时定义了不同级别数据的结构，发布了一系列的功能软件[20]，包括用于 FITS 格式文件操作的 FTOOLS；用于 FITS 格式文件读与写的 C 语言函数库 CFITSIO；用于天文 X 射线数据分析的软件 XANADU。

目前科学存档中心收录的数据主要有：雨燕伽玛射线天文台数据、钱德拉 X 射线天文台数据等。如康普顿伽玛射线天文台、伦琴 X 射线天文台等这些已经停止工作的卫星数据也可以在科学存档中心查询下载，更多探测设备信息可以在科学存档中心的网站㊷上查询。

科学存档中心新开发了数据检索系统 Xamin[21]，这是一种高效的数据检索系统。除了检索功能外，系统还提供了对用户表和查询记录保存以及其他功能。Xamin 不但可以查询科学存档中心的数据，还可以通过位置坐标以及时间对 Vizier 以及虚拟天文台的数据进行查询，并对科学存档中心的资源数据按照观测任务、波长、对象类别和名称进行分级。查询结果提供标准化接口、标记、不同输出格式、自定义误差等。目前，Xamin 提供浏览器的虚拟桌面版本和需要下载的命令行控制版本。官网有详尽的使用说明以及视频演示，方便用户快速熟悉查询流程，目前使用 Xamin 功能需要提交申请。

除了 Xamin 以外，旧版的浏览界面目前仍可以使用，数据查询平台可以查询到旧版搜索工具的使用方法。

5.1.2 红外科学存档中心

红外科学存档中心(Infrared Science Archive，IRSA)[22]收集美国宇航局红外和亚毫米波段观测的科学数据，其中包括许多巡天项目。它提供超过 200 亿个天体测量数据，涵盖全天 20 个波段。红外科学存档中心拥有强大的查询引擎用于访问数据库，同时支持虚拟天文台访问接口，并提供了红外天文数据分析和可视化工具。它同时隶属于美国宇航局与红外处理和分析中心(Infrared Processing and Analysis Center，IPAC)。

目前红外科学存档中心收录的数据主要来源有：斯皮策空间望远镜(Spitzer)、广域红外巡天探测者(WISE)、赫歇尔空间天文台(Herschel)、普朗克空间望远镜(Planck)、2 μm 全天巡视(2MASS)等。它还有一些其他时间比较久远的数据源，可以在官网的专题网站㊸上查询详细的数据来源。

红外科学存档中心推荐使用的搜索工具是 Atlas，Atlas 用来搜索红外科学存档中心的数据集，包括图像、光谱、光变曲线等，并将结果显示在网页上。在红外科学存档中心每个数据集都有简短的描述，巡天数据集配有全天合成图像。另外它支持包括 WISE、2MASS 在内的 15 个星表数据检索。

5.1.3 微波背景辐射数据存档中心

微波背景辐射数据存档中心(Legacy Archive for Microwave Background Data Analysis，LAMBDA)是高能天体物理科学存档中心的一部分，服务于专业从事微波背景辐射研究的人员，其网站㊹为从事微波背景辐射的研究人员提供数据、软件工具以及其他数据库的链接。目前它包含的数据集主要有：

(1)宇宙微波背景辐射各向异性实验，包含的项目有：Atacama 宇宙学望远镜(ACT)、普朗克项目(Planck)、微波背景辐射偏振观测计划(POLARBEAR)、Q/U 成像实验(QUIET)、南极点望远镜(SPT)、威尔金森微波各向异性探测器(WMAP)。

(2)宇宙微波背景辐射光谱实验，包含的项目有：宇宙学、天体物理和弥散辐射的绝对辐射计量(ARCADE)、宇宙背景探测器(COBE-FIRAS)。

(3)模拟数据，主要为宇宙微波背景模拟(CMB simulations)。

㊷ https://heasarc.gsfc.nasa.gov/
㊸ http://irsa.ipac.caltech.edu/frontpage/
㊹ http://lambda.gsfc.nasa.gov/

(4)前景(Foregrounds)⑮，包含的项目有：CHIPASS 1.4 GHz 连续星图、综合全天 H-alpha 波段星图、处理后 408 MH 全天射电星图、源于 Leiden/Dwingeloo 的全天 N(HI)图、94 GHz 预测尘埃星图、威尔金森微波各向异性探测器(Wilkinson Microwave Anisotropy Probe，WMAP)点源星表。

(5)红外，包含的项目有：宇宙背景探测器、亚毫米波天文卫星、红外天文卫星。

(6)微波背景辐射数据存档中心还包含多个宇宙微波背景辐射各向异性实验项目的数据。

5.1.4 Mikulski 太空望远镜存档中心

Mikulski 太空望远镜存档中心⑯是由美国宇航局资助的天文数据项目，负责收集光学、紫外和一部分近红外光谱数据，隶属于空间望远镜科学研究所(STScI)，它还是斯隆数字巡天的一部分，提供大型的搜索引擎，除了可以对单个项目搜索，还可以跨项目搜索，它包含的项目数据包括：ASTRO 天文台、数字化巡天项目、国际紫外探测器、哈勃太空望远镜等。相关数据可以通过网站⑰查询。

5.2 太阳物理学

目前美国宇航局管理和联合管理的太阳数据库为多个太阳虚拟天文台，这些虚拟天文台分别是：虚拟高能粒子天文台(VEPO)、虚拟格林威治天文台(VHO)、虚拟 ITM 天文台(VITMO)、虚拟磁层天文台(VMO)、虚拟辐射带天文台(ViRBO)、虚拟太阳天文台(VSO)、虚拟空间物理天文台(VSPO)、虚拟波天文台(VWO)、欧洲太阳观测网(EGSO)。其中虚拟太阳天文台是目前研究领域使用比较多的太阳数据库。

5.3 太阳系探索

美国宇航局对其历次的太阳系天体探索进行了归类，可以按照不同的天体对数据进行查询。这些天体分别是水星、金星、地球、月球、火星、木星、土星、天王星、海王星、矮行星、小行星以及彗星。目前美国宇航局以及合作机构已经发射了上百颗太阳系探测卫星，探测器产生的行星研究数据可以通过美国宇航局提供的网站⑱进行查询下载。

6 虚拟太阳天文台

6.1 简 介

虚拟太阳天文台(The Virtual Solar Observatory，VSO)⑲是一个把世界各地分散的太阳观测数据整合起来的数据库[23]，由哈佛-史密松天体物理中心、蒙大拿州立大学、美国国家太阳天文台、NASA/Goddard 太阳数据分析中心管理。虚拟太阳天文台采用分布式计算方式，当用户发送请求后，虚拟太阳天文台会将请求发送给数据提供方的数据目录，再从目录查询所需要的数据位置，最终从虚拟太阳天文台返回给用户，虚拟太阳天文台本身并不存放数据，所有服务可以通过访问虚拟太阳天文台的网站⑳使用。

研究者可以按照数据来源、观测物理量、维度、波段、时间分辨率、空间分辨率等对查找内容进行限定，以便精确查找所需数据[24]。还可以使用基于内容的搜索方式，通过从数据图像中计算出来的物理量查找数据。比如用某一空间区域的磁场高于某一阈值为条件，就可以查看其随时间的变化。方便的查询方式和大量数据的整合使得对太阳数据的统计研究成为可能。

虚拟太阳天文台数据可用于空间天气研究、日冕物质抛射预报、日震学、太阳表面下的活动预兆研究、太阳风起源、日冕加热机制研究、辐射变化、活动区特征研究等。

---

⑮ http://lambda.gsfc.nasa.gov/product/

⑯ http://archive.stsci.edu/

⑰ http://archive.stsci.edu/searches.html

⑱ http://pds.nasa.gov/

⑲ http://sdac.virtualsolar.org/cgi/search

⑳ http://vso.nso.edu/

虚拟太阳天文台的数据来源于各个与太阳观测有关的卫星以及地基观测点[51]，主要有 Hα 网络(HANET)，高山天文台(HAO)、SDO 联合科学操作中心(JSOC)、基彭霍尔太阳物理研究所(KIS)、大气与空间物理实验室(LASP)、空间与太阳物理实验室(LSSP)等。

6.2　查询方式

6.2.1　使用虚拟太阳天文台网页界面查询

(1)虚拟太阳天文台网页界面

虚拟太阳天文台网站可以实现按选择参数生成搜索条件，其中观测时间是必选参数，而物理观测量、数据来源、波段都是可选参数，支持多选。通过搜索功能，虚拟太阳天文台会直接链接到相关的数据提供方，将符合搜索条件的数据返回，并生成数据列表，数据的下载提供 .tar 压缩包下载和单个数据链接两种方式。

(2)虚拟太阳天文台命令行界面

虚拟太阳天文台提供了一个 Perl 脚本(query.pl)，可以从命令行下载数据。

(3)通过 IDL 下载

使用 IDL 的太阳软件包 SSWIDL，可以在命令行使用 vso_search 查找符合参数的数据，并由 vso_get 命令下载数据。用户还可以使用 VSO API 开发自己的界面。

6.2.2　生成视频

在虚拟太阳天文台网页界面上查询到所需数据后，可勾选所需数据(要求数据有可预览的图片或图标)，点击“添加至列表(Add to Shopping Cart)”按钮。如需继续添加其余数据，点击“继续添加至列表(Continue Adding to Cart)”。然后到列表中调整图片顺序，调整好后点击左下方“视频播放器(Movie Player)”按钮播放即可。

7　斯隆数字巡天

7.1　简　介

斯隆数字巡天(Sloan Digital Sky Survey，SDSS)[52]是利用美国新墨西哥州阿帕奇山上 2.5 m 光学望远镜进行的大规模巡天项目。从 2000 年开始运行，斯隆数字巡天获取了超过三百万个天体的多色测光和观测数据，覆盖了全天四分之一天区，创造了目前最为详尽的宇宙三维星图。

目前斯隆数字巡天最新发布的是第 12 批数据(DR12)，这也是 SDSS-III 最后一批数据，观测数据截止 2014 年 7 月。数据包含之前发布的使用重子振荡光谱巡天(Baryon Oscillation Spectroscopic Survey，BOSS)和阿帕奇波因特天文台星系演化实验(Apache Point Observatory Galactic Evolution Experiment，APOGEE)仪器观测的完整数据，还新增了通过 MARVELS 测量的恒星径向速度，DR12 的数据已经公布在网站[53]上。DR12 以及之前的版本数据都可以在斯隆数字巡天的数据库中查看和下载。目前斯隆数字巡天已经拥有超过 100 万星系、类星体、恒星的光学光谱数据，第 4 阶段工作也已经开始，预计为期 6 年。

7.2　查　询

斯隆数字巡天每次数据发布包含 4 种类型的数据[25]：图像、可见光光谱、红外光谱、星表数据，其中，星表数据是从图像和光谱数据中获取的亮度、红移等数据。斯隆数字巡天为这些数据提供了不同的查询方式，包括：目标的网页查询、交互式图像和光谱查询科学数据归档服务(Science Archive Server)、基于浏览器的访问星表归档服务数据库(SkyServer)、基于 SQL 的查询界面(CasJobs 需要注册后才能使用)，可以直接下载 DR12 的 FITS 格式文件(DR12 FITS)，另外还可以查看数据归档服务

---

[51] http://docs.virtualsolar.org/

[52] http://www.sdss.org/

[53] http://www.sdss.org/dr12/

的目录结构、文件格式以及每个文件内容(Data Model)，Data Model 访问数据有两种方式，一种是通过数据模型的目录结构，另一种是通过数据模型的文件索引。

8 澳大利亚国家望远镜中心

8.1 简 介

澳大利亚国家望远镜中心(Australia Telescope National Facility，ATNF)隶属于澳大利亚国立科研机构——英联邦科学和工业研究组织(Commonwealth Scientific and Industrial Research Organisation，CSIRO)，由 CSIRO 的天文与空间科学部门运行管理。目前 ATNF 的大部分观测数据已实现开放共享，通过 ATNF 网站㊹可实现数据检索。

ATNF 有 4 个主要观测设备，最重要的是位于帕克斯的 64 m 射电望远镜(Parkes 64-metre Radio Telescope)[26]。64 m 射电望远镜 1961 年开始服役，经过了多次升级改造，以适应多波束焦平面阵列接收设备，在国际射电望远镜排名中位于世界前列。该望远镜为 64 m 全可动碟型望远镜，是南半球第二大射电望远镜，也是世界上第一批大型碟型天线之一。目前已知的 2 000 多颗脉冲星超过一半以上是它发现的，64 m 射电望远镜是少数几个接收过阿波罗 11 号登月电视直播信号的射电天线，被誉为澳大利亚最成功的科学仪器。时至今日，64 m 射电望远镜仍参与了众多探测器的数据回收任务，包括水手 2 号、水手 4 号、旅行者号、乔托号、伽利略探测器和卡西尼-惠更斯探测器等。

澳大利亚 SKA 探路者射电望远镜(Australian Square Kilometre Array Pathfinder，ASKAP)及 MWA (Murchison Widefield Array)都属于 CSIRO。前者由 36 台 12 m 直径的天线组成并协同工作，后者由 2 048 面低频天线组成，数据由 ICRAR 负责管理与维护[27-28]。

除了独立观测，ATNF 的射电望远镜还可与澳大利亚和新西兰的其他射电天线一起组成长基线。

8.2 ATNF 的脉冲星表介绍

8.2.1 ATNF 脉冲星表简介

ATNF 脉冲星表[29]包括了所有已发表的自转供能脉冲星，星表中包括反常 X 射线脉冲星(Anomalous X-ray Pulsar，AXP)以及软伽马射线复现源(Soft γ-ray Repeater，SGR)，但不包括吸积供能的脉冲星以及今年发现的 X 射线毫秒脉冲星，ATNF 脉冲星表可以通过网站㊺在线查询。

8.2.2 数据获取

ATNF 脉冲星表可以通过网页在线查询，查询输入项为脉冲星相关参数，附加参考信息等。数据查询的输出可以选择多种预设的表格输出模式，同时提供 68 个预设参数，并可以自定义新的参数。查询结果支持绘图输出，用户可以将参数的分布情况绘制为二维图像，复合参数可通过数学操作符或函数定义，也可以通过脉冲星的通用名称实现数据查询。

ATNF 还为专业研究者提供了专用的查询页面㊻，其中提供的预设参数增加了 98 个。用户除了使用网页查询相关数据，还可以通过 unix/linux 系统的命令行直接获取 .tar 文件数据。

8.3 ATNF 数据中心

CSIRO 提供未经处理的 ATNF 射电天文数据下载。数据来自澳大利亚望远镜致密阵(Telescope Compact Array，ATCA)、帕克斯射电望远镜(Parkes Radio Telescope)、莫普拉天文台(Mopra Observatory)以及长基线望远镜阵(Long Baseline Array，LBA)。数据可从澳大利亚望远镜在线数据库(Telescope Online Archive，ATOA)和 CSIRO 数据访问门户网站的搜索表格得到。

而对于澳大利亚平方千米阵列探索计划(Australian Square Kilometre Array Pathfinder，ASKAP)[30]的数据，CSIRO 的 ASKAP 科学数据库(CSIRO ASKAP Science Data Archive，CASDA)提供经过处理的

㊹ http://www.atnf.csiro.au/
㊺ http://www.atnf.csiro.au/research/pulsar/psrcat/
㊻ http://www.atnf.csiro.au/observers/data/index.html

数据，包括校准了能见度的连续谱数据、图像和图像立方体以及探测目录。这些数据可由 CSIRO 数据访问门户网站或者虚拟天文台服务获取。

9　美国国家射电天文台

9.1　简　介

美国国家射电天文台(National Radio Astronomy Observatory，NRAO)[57]在 2003 年发布了一个数据库系统，提供网页门户[58]，并指导用户如何获取数据。这一数据库系统允许用户通过网页提交数据请求。2008 年，数据库系统已支持允许用户检索得到下列公开数据：

(1)所有甚大阵(Very Large Array，VLA)的数据及图像；

(2)所有甚长基线阵(Very Long Baseline Array，VLBA)自 1997 年以来的数据和图像；

(3)绿岸望远镜(The Robert C. Byrd Green Bank Telescope，GBT)数据。

所有从美国国家射电天文台望远镜得到的数据都会在这个系统中记录并做成列表，可以直接通过 FTP 下载。数据的保护周期一般为 1 年，对于 12 个月以前的数据，数据库无限制开放共享。

9.2　数据获取

美国国家射电天文台数据库每日收录甚大阵、甚长基线阵及绿岸望远镜的新数据[31]。用户可通过网页客户端进行数据检索，同时也可以通过 FTP 方式获取公开数据和私有数据。基础搜索可通过观测项目 ID、观测者姓名、观测日期(时段)等进行搜索。高级搜索涵盖了数据类型、射电源名称、射电源坐标、灵敏度、分辨率以及频率范围等信息。

而对于绿岸望远镜观测一年内的私有数据下载时一般需要先得到一个密码。在数据私有阶段，用户需要通过密码下载数据。这一密码可以通过致电或者邮件询问美国国家射电天文台数据管理人员得到。当提交数据请求后，公共领域的数据会被写入一个默认的 FTP 目录中。私有数据则会被复制到一个临时保密子目录。

9.3　Data Vault

2007 年美国国家射电天文台发布了一个应用 Data Vault，用以简化访问美国国家射电天文台数据库的流程。该应用允许用户使用自由文本检索、浏览、下载和查看相关数据的详细信息。Data Vault 由 web.py 建立的控制器、谷歌网页工具包(GWT)、浏览器、MySQL 数据库等组成，应用 web.py 实现模型-视图-控制器(Model-View-Controller，MVC)设计。Data Vault 还使用了一个 Ajax 客户端，支持链接数据到其他网页工具或服务插件。Data Vault 提供给用户一个美国国家射电天文台为中心的数据库，并可以链接或提供尽可能丰富的信息。

在 Data Vault 之前，绿岸望远镜数据库与甚大阵以及甚长基线阵数据库系统是完全独立的。在 2008 年之后，Data Vault 成为一个链接多个数据库的一站式入口，不仅可以链接到上述 3 个数据库，也可在必要时加入其他数据库。未来还将添加虚拟天文台搜索服务，让研究人员可以无缝链接访问其他在线数据库。

9.4　绿岸望远镜数据

绿岸望远镜是世界上首台可以观测米波到毫米波段(3.0 m~3 mm)的单天线射电望远镜。为 100 m × 110 m 的椭圆形天线，观测范围可以覆盖 85%的天区。

2011 年起，绿岸望远镜的全部数据归档到美国国家射电天文台数据库中，利用数据存档访问工具(NRAO-wide archive access tool，AAT)进行检索，2012 年 10 月开始公开发布绿岸望远镜观测数据。目前归档到美国国家射电天文台的绿岸望远镜数据仅包含脉冲星的到达时间数据，绿岸望远镜有两种数据格式，一种是原始 FITS 数据文件，一种是单天线 FITS 数据文件(SDFITS)。

---

[57] https://www.nrao.edu/

[58] http://archive.nrao.edu/

10 总 结

系统回顾了中国各天文台站的数据归档与检索相关情况，综述了中国科学院下属各天文机构的数据情况，对于其他研究机构或大学的数据未做统计。我国各天文台站中，国家天文台的数据规模最为庞大，并负责中国天文数据中心的建设、数据管理与维护工作。其他台站以各自的观测设备为基础分别建设了相应的数据发布平台，并实现了数据的在线检索服务。国家天文台、新疆天文台的部分数据已注册到国际虚拟天文台联盟，可以通过虚拟天文台相关工具实现在线检索。

针对国外天文机构数据管理与发布问题，只涵盖了斯特拉斯堡天文数据中心、美国宇航局空间科学数据中心、欧南台科学数据中心、虚拟太阳天文台、斯隆数字巡天、澳大利亚国家望远镜中心、美国国家射电天文台的部分设备及数据情况。国外大部分天文机构的数据发布环境基于虚拟天文台的相关标准、协议建设，在统一的协议标准下实现检索为科研人员获取数据提供了便利。

在整理的过程中未考虑各天文研究机构的网络资源与计算资源。大口径观测设备及阵列望远镜的诞生，使得海量天文数据不得不依赖于高性能计算系统管理，事实上各天文机构均有各自的高性能计算环境。高性能计算设备、存储设备的采用大大提高了海量天文数据的检索与预处理速度。

参考文献：

[1] 崔辰州，赵永恒. 中国虚拟天文台体系结构 [J]. 天文研究与技术——国家天文台台刊，2004，1(2)：140-151.
Cui Chenzhou，Zhao Yongheng. Architecture of Chinese Virtual Observatory [J]. Astronomical Research & Technology——Publications of National Astronomical Observatories of China，2004，1(2)：140-151.

[2] 钟守波，韩波，张彦霞，等. 天文大数据管理工具的设计与实现 [J]. 天文研究与技术，2015，12(4)：510-517.
Zhong Shoubo，Han Bo，Zhang Yanxia，et al. Design and implementation of a software tool package for managing massive astronomical data [J]. Astronomical Research & Technology，2015，12(4)：510-517.

[3] 李建，崔辰州，何勃亮，等. 天文数据库回顾与展望 [J]. 天文学进展，2013，31(1)：1-16.
Li Jian，Cui Chenzhou，He Boliang，et al. Review and prospect of the astronomical database [J]. Progress in Astronomy，2013，31(1)：1-16.

[4] 李文. 虚拟天文台环境下的海量数据存储与访问技术研究 [D]. 天津：天津大学，2007.

[5] 桑健，赵永恒，崔辰州. 中国虚拟天文台数据访问服务 [J]. 天文研究与技术——国家天文台台刊，2004，1(3)：216-228.
Sang Jian，Zhao Yongheng，Cui Chenzhou. Data access service of China-VO [J]. Astronomical Research & Technology——Publications of National Astronomical Observatories of China，2004，1(3)：216-228.

[6] Zhao Gang，Zhao Yongheng，Chu Yaoquan，et al. LAMOST spectral survey—an overview [J]. Research in Astronomy and Astrophysics，2012，12(7)：723-734.

[7] Luo Ali，Zhang Haotong，Zhao Yongheng，et al. Data release of the LAMOST pilot survey [J]. Research in Astronomy and Astrophysics，2012，12(9)：1243-1246.

[8] Lee Y S，Beers T C，Carlin J L，et al. Application of the SEGUE Stellar Parameter Pipeline to LAMOST Stellar Spectra [J]. The Astronomical Journal，2015，150(6)：187-204.

[9] Ma Bin, Shang Zhaohui, Wang Lifan, et al. The test of the 10k×10k CCD for Antarctic Survey Telescopes (AST3) [C] // Proceedings of SPIE. 2012.

[10] Li Zhengyang, Yuan Xiangyan, Cui Xiangqun, et al. Status of the first Antarctic Survey Telescopes for Dome A [C] // Proceedings of SPIE. 2012.

[11] Zou Hu, Jiang Zhaoji, Zhou Xu, et al. South Galactic Cap u-band Sky Survey (SCUSS): data reduction [J]. The Astronomical Journal, 2015, 150(4): 104-120.

[12] 刘健，林钢华. 构建虚拟太阳天文台的中国数据结点 [J]. 天文研究与技术——国家天文台台刊, 2008, 5(3): 241-247.
Liu Jian, Lin Ganghua. Construction of the Chinese data-node system for virtual solar observatory [J]. Astronomical Research & Technology——Publications of National Astronomical Observatories of China, 2008, 5(3): 241-247.

[13] 逯登荣，孙继先，刘梁，等. 毫米波射电天文数据库 [J]. 科研信息化技术与应用, 2011, 1(3): 81-86.
Lu Dengrong, Sun Jixian, Liu Liang, et al. Millimeter-Wave radio astronomy database [J]. e-Science Technology & Application, 2011, 2(3): 81-86.

[14] 袁瑾，苟伟，赵融冰，等. 65 米射电望远镜中频传输分配系统的研制 [J]. 天文研究与技术, 2015, 12(3): 262-269.
Yuan Jin, Gou Wei, Zhao Rongbing, et al. Development of an IF-signal transmission and distribution system for the 65m telescope of the SAO [J]. Astronomical Research & Technology, 2015, 12(3): 262-269.

[15] Liu Yingbo, Wang Feng, Ji Kaifan, et al. NVST data archiving system based on Fastbit NoSQL database [J]. Journal of the Korean Astronomical Society, 2014, 47(3): 115-122.

[16] Perret E, Boch T, Bonnarel F, et al. Working together at CDS: the symbiosis between astronomers, documentalists, and IT specialists [C] // Astronomical Society of the PacificConference Series. 2014.

[17] Delmotte N, Dolensky M, Micol A, et al. The 2006 ESO science archive survey [J]. The Messenger, 2006, 125: 41-43.

[18] Pence W D, Chiappetti L, Page C G, et al. Definition of the Flexible Image Transport System (FITS), version 3.0 [J]. Astronomy & Astrophysics, 2010, 524: A42-A81.

[19] Thomas B, Jenness T, Economou F, et al. Significant problems in FITS limit its use in modern astronomical research [C] // Astronomical Data Analysis Software and Systems XXIII. 2014.

[20] Drake S A, Smale A P, McGlynn T A, et al. The HEASARC in 2013 and Beyond: NuSTAR, Astro-H, NICER [C] // American Astronomical Society, HEAD Meeting. 2013.

[21] Barrett P. HEASARC—The High Energy Astrophysics Science Archive Research Center [C] // Cataclysmic Variables and Related Physics, 2nd Technion Haifa Conference. 1993.

[22] Berriman G B. The NASA/IPAC Infrared Science Archive (IRSA) as a resource in supporting observatory operations [C] // Proceedings of SPIE. 2008.

[23] Hill F. The Virtual Solar Observatory concept [C] // Proceedings of the 1st Solar and Space Weather Euroconference. 2000: 569-574.

[24] Davey A R. The Virtual Solar Observatory: new data, big data and better queries [C] // American Astronomical Society, AAS Meeting. 2014.

[25] Eisenstein D J, Weinberg D H, Agol E, et al. SDSS-III: massive spectroscopic surveys of the distant universe, the Milky Way, and extra-solar planetary systems [J]. The Astronomical Journal, 2011, 142(3): 72-95.

[26] Lyne A G, Manchester R N, Lorimer D R, et al. The Parkes southern pulsar survey-II. final results and population analysis [J]. Monthly Notices of the Royal Astronomical Society, 1998, 295(4): 743-755.

[27] Lonsdale C J, Cappallo R J, Morales M F, et al. The Murchison Widefield Array: design overview [J]. Proceedings of the IEEE, 2009, 97(8): 1497-1506.

[28] Tingay S J, Goeke R, Bowman J D, et al. The Murchison Widefield Array: the Square Kilometre Array Precursor at low radio frequencies [J]. Publications of the Astronomical Society of Australia, 2013, 30: 7-27.

[29] Manchester R N, Hobbs G B, Teoh A, et al. The Australia Telescope National Facility pulsar catalogue [J]. The Astronomical Journal, 2005, 129(4): 1993-2006.

[30] Johnston S, Taylor R, Bailes M, et al. Science with ASKAP [J]. Experimental Astronomy, 2008, 22(3): 151-273.

[31] DuPlain R, Benson J, Sessoms E. Data Vault: providing simple web access to NRAO data archives [C] // Proceedings of SPIE. 2008.

Astronomical Data Query and Release Review

Zhang Hailong[1,2], Ye Xinchen[1], Li Huijuan[1], Wang Jie[1], Wang Wanqiong[1], Tohtonur[1], Nie Jun[1], Cui Chenzhou[3], Liu Liang[4], Chen Junyi[5], Chen Xiao[6], Xue Yansong[6], He Boliang[3], Li Changhua[3], Zhao Qing[8], Xiao Jian[9], Fan Dongwei[3], Cao Zihuang[7], Li Shanshan[3], Mi Linying[3], Yang Zherui[4]

(1. Xinjiang Astronomical Observatory, Chinese Academy of Sciences, Urumqi 830011, China, Email: zhanghailong@xao.ac.cn;
2. Key Laboratory of Radio Astronomy, Chinese Academy of Sciences, Nanjing 210008, China; 3. National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China; 4. Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210008, China; 5. Yunnan Observatories, Chinese Academy of Sciences, Kunming 650011, China;
6. Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China;
7. LAMOST, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China;
8. Tianjin University of Science and Technology, Tianjin 300222, China;
9. Tianjin University, Tianjin 300072, China)

Abstract: Astronomical big data becomes the challenges for astronomy research. The challenges include how to archive, query and manage the massive observational and processed data efficiently and timely, how to accelerate the output of astronomy research. Analysis and management of the big astronomical data based on information technologies will help the astronomers to understand our universe. The Virtual Observatory (VO) provides protocols and standards on linking global astronomical data seamlessly and transparently, and VO protocols normalized publish and query methods of astronomical data. This paper summarizes astronomical data release and retrieval of global observation facilities from mainstream observatories.

Key words: Observational data; Data center; Virtual Observatory; Data archiving; Data release